

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/112780/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Pitt, Daniel ORCID: <https://orcid.org/0000-0002-8429-2957>, Sevane Fernandez, Natalia ORCID: <https://orcid.org/0000-0003-4766-6291>, Nicolazzi, Ezequiel L., MacHugh, David E., Park, Stepehn D. E., Colli, Licia, Martinez, Rodrigo, Bruford, Michael W. ORCID: <https://orcid.org/0000-0001-6357-6080> and Orozco Ter Wengel, Pablo ORCID: <https://orcid.org/0000-0002-7951-4148> 2019. Domestication of cattle: two or three events? Evolutionary Applications 12 (1) , pp. 123-136. 10.1111/eva.12674 file

Publishers page: <http://dx.doi.org/10.1111/eva.12674>  
<<http://dx.doi.org/10.1111/eva.12674>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.





See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



## ORIGINAL ARTICLE

## Domestication of cattle: Two or three events?

Daniel Pitt<sup>1</sup>  | Natalia Sevane<sup>1</sup>  | Ezequiel L. Nicolazzi<sup>2</sup> | David E. MacHugh<sup>3,4</sup> | Stephen D. E. Park<sup>5</sup> | Licia Colli<sup>6</sup> | Rodrigo Martinez<sup>7</sup> | Michael W. Bruford<sup>1</sup>  | Pablo Orozco-terWengel<sup>1</sup> 

<sup>1</sup>School of Biosciences, Cardiff University, Cardiff, UK

<sup>2</sup>Parco Tecnologico Padano (PTP), Lodi, Italy

<sup>3</sup>Animal Genomics Laboratory, UCD School of Agriculture and Food Science, UCD College of Health and Agricultural Sciences, University College Dublin, Dublin, Ireland

<sup>4</sup>UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland

<sup>5</sup>IdentIGEN Ltd., Dublin, Ireland

<sup>6</sup>Istituto di Zootecnica e BioDNA Centro di Ricerca sulla Biodiversità e sul DNA Antico, Università Cattolica del S. Cuore di Piacenza, Piacenza, Italy

<sup>7</sup>Corporación Colombiana De Investigación Agropecuaria (Corpoica), Centro de investigaciones Tibaitatá, Bogotá, Colombia

## Correspondence

Pablo Orozco-terWengel, School of Biosciences, Cardiff University, Cardiff, CF10 3AX, UK.  
Email: orozco-terwengelpa@cardiff.ac.uk

## Funding information

BBSRC SWBio DTP PhD studentship, Grant/Award Number: BB/M009122/1; FACCE-JPI ERA-NET Smart Agriculture Program, Grant/Award Number: BB/M019276/1; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: DLV-655100

## Abstract

Cattle have been invaluable for the transition of human society from nomadic hunter-gatherers to sedentary farming communities throughout much of Europe, Asia and Africa since the earliest domestication of cattle more than 10,000 years ago. Although current understanding of relationships among ancestral populations remains limited, domestication of cattle is thought to have occurred on two or three occasions, giving rise to the taurine (*Bos taurus*) and indicine (*Bos indicus*) species that share the aurochs (*Bos primigenius*) as common ancestor ~250,000 years ago. Indicine and taurine cattle were domesticated in the Indus Valley and Fertile Crescent, respectively; however, an additional domestication event for taurine in the Western Desert of Egypt has also been proposed. We analysed medium density Illumina Bovine SNP array (~54,000 loci) data across 3,196 individuals, representing 180 taurine and indicine populations to investigate population structure within and between populations, and domestication and demographic dynamics using approximate Bayesian computation (ABC). Comparative analyses between scenarios modelling two and three domestication events consistently favour a model with only two episodes and suggest that the additional genetic variation component usually detected in African taurine cattle may be explained by hybridization with local aurochs in Africa after the domestication of taurine cattle in the Fertile Crescent. African indicine cattle exhibit high levels of shared genetic variation with Asian indicine cattle due to their recent divergence and with African taurine cattle through relatively recent gene flow. Scenarios with unidirectional or bidirectional migratory events between European taurine and Asian indicine cattle are also plausible, although further studies are needed to disentangle the complex human-mediated dispersion patterns of domestic cattle. This study therefore helps to clarify the effect of past demographic history on the genetic variation of modern cattle, providing a basis for further analyses exploring alternative migratory routes for early domestic populations.

## KEYWORDS

approximate Bayesian computation, *Bos indicus*, *Bos Taurus*, demographic modeling, domestication history, SNP array

## 1 | INTRODUCTION

Between the late Pleistocene and early Holocene, the most commonly occurring cattle species was the aurochs (*Bos primigenius*), ranging from northern Africa to both the Atlantic and Pacific coasts of Eurasia (Zeuner, 1963). This formerly widespread wild species recently became extinct, with the last recorded herd found in 1627 AD in Poland (Götherström et al., 2005). Similar to sheep and goats, archaeological and genomic evidence suggests that the ancestors of taurine cattle (*Bos taurus*) were domesticated from *Bos primigenius* in the Fertile Crescent during the Neolithic, more than 10,000 years ago (YA; Bruford, Bradley, & Luikart, 2003; Ajmone-Marsan, Garcia, & Lenstra, 2010; MacHugh, Larson, & Orlando, 2017). However, approximately 1,500 years later a second domestication event took place in the Indus Valley from *Bos primigenius nomadicus*, separated from the taurine branch ~250–330,000 YA, eventually giving rise to the extant indicine cattle (*Bos indicus*), often also termed zebu cattle (Loftus, MacHugh, Bradley, Sharp, & Cunningham, 1994).

Despite support for the extant distribution of cattle arising from two main domesticated lineages, a third domestication event has been hypothesized to have occurred in north-east Africa about 8,000–9,000 YA, giving rise to the divergent African taurine cattle. Support for this hypothesis derives from archaeozoological evidence and genetic data from contemporary cattle and aurochs. The archaeozoological evidence is supported by comparative osteological analyses between ancient wild and domestic cattle from Europe, Asia and Africa (Applegate, Gautier, & Duncan, 2001; Grigson, 1991, 2000; Stock & Gifford-Gonzalez, 2013). Bradley, MacHugh, Cunningham, and Loftus (1996), using maternal mitochondrial DNA (mtDNA), showed that African cattle feature a higher frequency of the T1 mitochondrial haplogroup than is common in other regions, estimated that the separation between African and European taurine ancestors occurred 22,000–26,000 YA (earlier than the Fertile Crescent domestication), and found patterns of population expansions consistent with domestication that were more recent than the corresponding signature of African/European divergence. These results are supported by analyses of extant taurine cattle from Europe, the Middle East and Africa, as well as extinct British aurochs (Troy et al., 2001). Furthermore, evidence from nuclear DNA analyses (Bovine Hapmap Consortium et al., 2009; Hanotte et al., 2002; Pérez-Pardal et al., 2010) is consistent with a local domestication of African taurine cattle and the subsequent admixture of Near East and the Indus Valley cattle in Africa. However, Bonfiglio et al. (2012) support the Near Eastern origin of the T1 mitochondrial DNA haplogroups and defended that the North African subhaplogroup T1d could have originated in already domesticated cattle shortly after their arrival from the Near East. Along this line, analyses on whole-genome SNP arrays supported two domestication events for taurine and indicine lineages followed by introgression from wild aurochs in Africa, East Asia and Europe (Decker et al., 2014). Taken together, controversial archaeological and genetic data seem to support both two and three domestication events in cattle; therefore, an analysis

using genealogical modelling of alternative scenarios that could give rise to the extant patterns of cattle genetic diversity needs to be addressed.

Taurine cattle dispersed quickly after their domestication north-west from the Fertile Crescent through Turkey into the Balkans and into northern Italy, either following a Mediterranean coastline or a route partially along the Danube River, and subsequently dispersing across Europe (Figure 1) (Beja-Pereira et al., 2006; Pellecchia et al., 2007). Taurine cattle may have also migrated along the northern coast of Africa, eventually crossing into the Iberian Peninsula and admixing with local cattle (Figure 1) (Beja-Pereira et al., 2006; Decker et al., 2009). Similarly, indicine cattle also dispersed far beyond their domestication centre in the Indus Valley, reaching China and much of South-East Asia (Figure 1) (Ajmone-Marsan et al., 2010). Additionally, a modern indicine migration via pastoralists into eastern Africa (~2,500–3,500 YA) and subsequently throughout central and southern areas of the continent has been described (Figure 1) (Decker et al., 2014; Hanotte et al., 2002; Payne & Hodges, 1997).

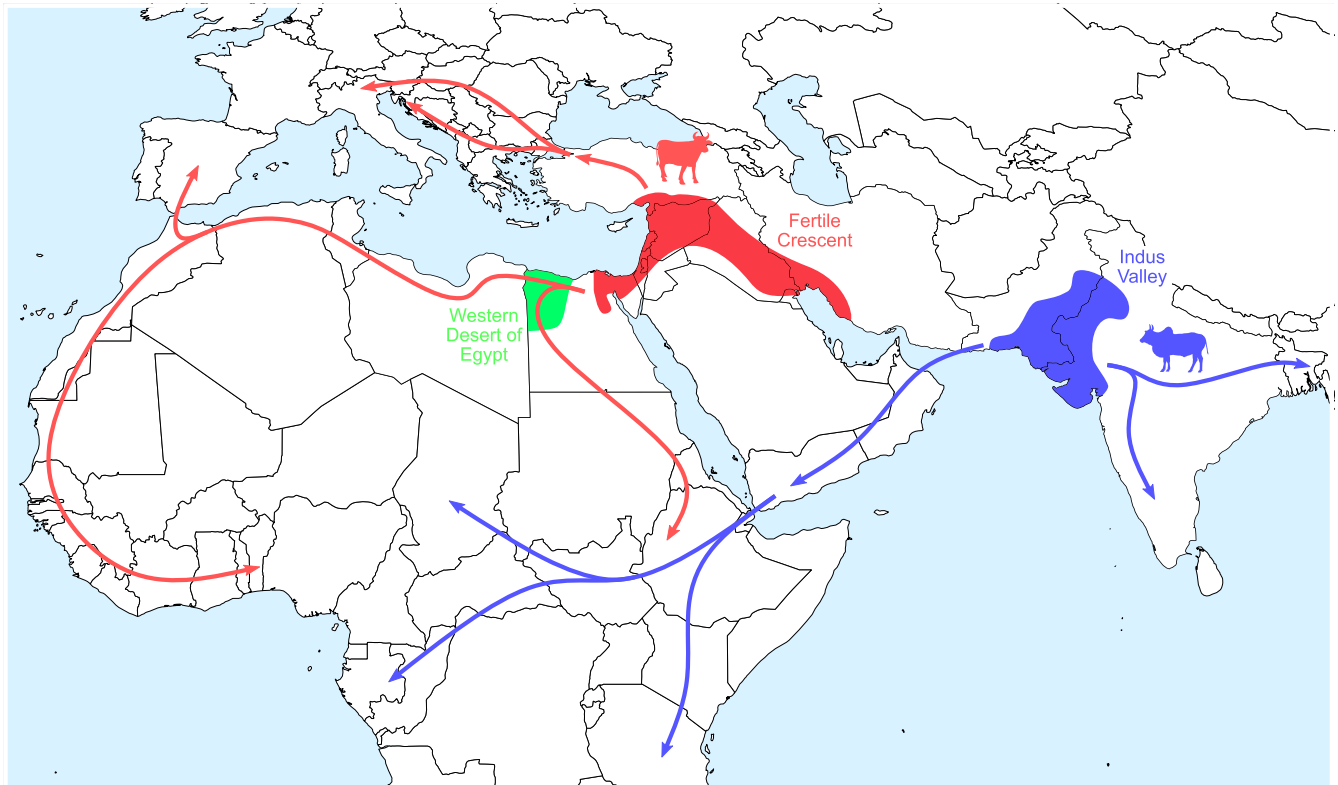
Previous studies using traditional molecular markers (e.g., mitochondrial DNA and microsatellites) generated insights into the genetics underlying the domestication history and postdomestication processes, such as the taurine dispersal patterns in Europe (Beja-Pereira et al., 2006; Bradley et al., 1996; MacHugh, Shriver, Loftus, Cunningham, & Bradley, 1997). However, the development of newer genomic technologies, such as SNP arrays, has enabled the rapid collection of information for thousands of markers with high precision, facilitating comprehensive surveys of genomewide variation in domestic cattle (Bovine Hapmap et al. 2009).

Here, we collated a data set of ~54,000 SNPs genotyped in 180 cattle populations from across the world, representing both indicine and taurine breeds (Supporting Information Table S1), to characterize the genetic diversity, population structure and demographic history of extant cattle. To disentangle the long-standing debate around the existence of a third domestication event in north-east Africa, we used approximate Bayesian computation to model a variety of possible scenarios that could give rise to the extant patterns of cattle genetic diversity, including analyses to determine whether modern cattle derive from two or three domestication events and migratory patterns among breeds.

## 2 | MATERIALS AND METHODS

### 2.1 | SNP array data

Illumina's BovineSNP50 v.1, v.2 and Bovine High Density BeadChip 770k SNP array data were merged from several previously published studies (Supporting Information Table S1; Bovine Hapmap et al. 2009; Decker et al., 2009; Gautier, Laloë, & Moazami-Goudarzi, 2010; McTavish, Decker, Schnabel, Taylor, & Hillis, 2013; Decker et al., 2014; Mbole-Kariuki et al., 2014; Orozco-terWengel et al., 2015; Park et al., 2015; Iso-Touru et al., 2016; Upadhyay et al., 2017; Pitt et al., 2018). After removing duplicated copies of individuals, the



**FIGURE 1** Hypothesized main domestication sites and migration routes of taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle, including the postulated third domestication site in Egypt. Lesser migration routes, such as the dispersal across Europe, are not depicted

final data set comprised ~54,000 SNPs in most samples and corresponded to 3,196 individuals representing 180 breeds/populations of both taurine ( $n = 2,041$ ) and indicine ( $n = 408$ ) cattle, including hybrids of the two subspecies (Supporting Information Table S1). Additionally, one individual was genotyped from the archaeological remains of a *B. p. primigenius* animal from England (Supporting Information Table S1; Park et al., 2015) and used as an outgroup. The UMD3.1 bovine assembly was used as a reference to map the genomic positions for each SNP. All data sets were merged and quality controlled using PLINK 1.90 (Chang et al., 2015; Purcell et al., 2007). Across all breeds, only autosomal SNPs were retained and any locus with a minor allele frequency less than 5% and a call rate less than 90% was removed, leaving 8,081 SNPs (~8 k) for further analyses.

## 2.2 | Genetic variation and population divergence

The molecular inbreeding coefficient ( $F_{IS}$ ), and observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosities per breed were calculated using PLINK on the breeds with a sample size of at least eight individuals (breeds with lower sample sizes were not included to avoid potentially biased estimates deriving from unrepresentative sample sizes). Welch's  $t$  test was used to test for differences between  $H_e$  and  $H_o$  within each breed. Admixture v1.3 (Alexander, Novembre, & Lange, 2009) was used to assess population structure in the data for values of  $K$  (the number of clusters in the data) between 1 and 150.

The predictive accuracy for each  $K$  was estimated using the fivefold cross-validation (CV) procedure implemented in Admixture. For this analysis, we further pruned the 8K data set for linkage disequilibrium (LD) in PLINK, using a sliding window of 50 kb, a step size of 10 SNPs, and randomly removing one SNP from each pairwise comparison with an  $r^2$  of 5% or higher. Sixty-one values of  $K$  ranging between 1 and 150 were tested. Due to the computational power required for assessing higher values of  $K$ , we restricted the analysis to all values of  $K$  between 1 and 50, and used intervals of 5  $K$  from 55 to 100, and examined  $K = 150$ . The partition solutions were visualized using the POPHELPER R package (Francis, 2017).

Multidimensional scaling (MDS) was calculated in PLINK using Hamming distances to determine the distances between individuals and breeds across 20 dimensions. The first two major axes explaining the highest proportion of the variance in the data set were visualized using R (R Core Team, 2014).  $F_{ST}$  was also estimated between all pairs of populations featuring more than one sample, and the resultant matrix was used to construct a neighbour-net tree in SplitsTree v.4.14.4 (Huson & Bryant, 2006).

## 2.3 | Approximate Bayesian computation (ABC)

The domestication history of cattle was analysed with approximate Bayesian computation (ABC) as implemented in the software ABCtoolbox (Wegmann, Leuenberger, Neuenschwander, & Excoffier, 2010) with simulations generated using Fastsimcoal2 (Excoffier,



Dupanloup, Huerta-Sánchez, Sousa, & Foll, 2013; Excoffier & Foll, 2011). To reduce the computational power required to perform the analyses, as well as to reduce the effect of ascertainment bias (e.g., Barbato et al., 2017; Kijas et al., 2012; Orozco-terWengel et al., 2015), the 8K data set was pruned in PLINK as above but with an  $r^2$  of 2.5%, resulting in 2,202 SNPs (2K). Four breeds were selected for these analyses, a European taurine (TaurEU; Normande [NOR]), an African taurine (TaurAF; Somba [SOM]), an African indicine (ZebuAF; Madagascar Zebu [ZMA]) and an Asian indicine (ZebuAS; Kankraj [KAN]) (Table 1). These breeds were chosen avoiding outlier breeds based on the MDS and neighbour-net analyses. Two additional breed sets were selected with the same criteria and used as replicates for the ABC analysis to assess that the results obtained were not dependent on the breed set used but rather a general outcome reflective of the domestication history of cattle (Table 1). To ensure the 2K data set retained information about the genealogical history of these breeds, we compared the results of running the Admixture analysis on each set of four breeds using the 2K and 8K data sets. The 2K data set generated qualitatively the same results as the 8K data set. The observed data for each breed set were analysed with Arlequin 3.5 (Excoffier & Lischer, 2010) to produce 34 summary statistics describing the genetic variation within and among breeds (Supporting Information Table S3).

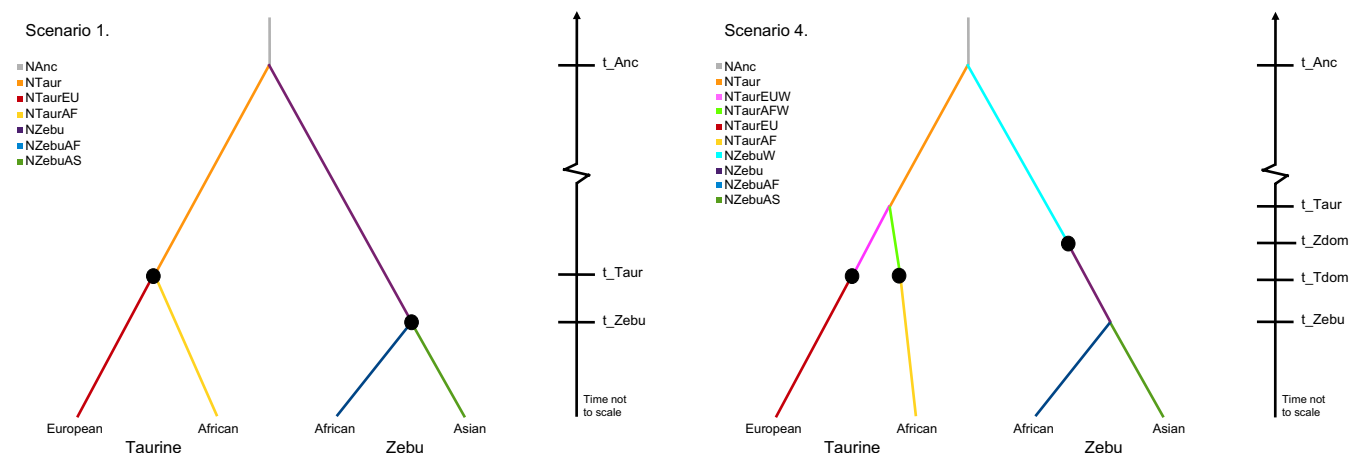
The underlying model used in the ABC analyses for the four geographically and genetically divergent groups assumed that: (a)

ZebuAF and ZebuAS have a common indicine ancestor; (b) TaurEU and TaurAF have a common taurine ancestor; (c) indicine and taurine have a common aurochs ancestor; and (d) bifurcation of aurochs occurred approximately 250,000 YA (Bovine Hapmap et al. 2009; Bradley et al., 1996). Initially, four alternative scenarios were tested to discriminate among the main demographic histories involving two or three domestication events, modelled as bottlenecks, and migratory patterns among breeds (examples shown in Figure 2; full detail in Supporting Information Figure S1). The software ABCtoolbox was used to parametrize the models to be simulated and to control Fastsimcoal2 for the random generation of the 1 million reverse coalescent simulations drawing parameter values from defined prior distribution ranges for each scenario (Supporting Information Table S2). Based on a set of 5,000 randomly selected simulations, Spearman's rank correlations between the 34 summary statistics were calculated to test for the presence of correlations between summary statistics. After Bonferroni correction, summary statistics with consistent significant negative or positive correlations to others were removed from further analyses, reducing the set of statistics for the ABC analysis from 34 to 17. The fitting of the simulations to the observed data was further determined by identifying if each of the summary statistics for the observed data fell within the 95% quantiles of the distribution of the corresponding summary statistics calculated on the simulated data for each scenario. Posterior estimates for each parameter in each scenario were obtained with

**TABLE 1** Breeds sets used on the modelling of domestication history of cattle with approximate Bayesian computation (ABC)

Breed set <sup>a</sup>	TaurEU	TaurAF	ZebuAF	ZebuAS
1	Normande (NOR)	Somba (SOM)	Zebu from Madagascar (ZMA)	Kankraj (KAN)
2	Abondance (ABO)	N'dama (NDA)	Zebu Fulani (ZFU)	Bhagnari (BAG)
3	Montbeliard (MON)	Baoule (BAO)	Zebu Bororo (ZBO)	Dajal (DAJ)

<sup>a</sup>A breed in each group was selected to represent European taurine (TaurEU), African taurine (TaurAF), African indicine (ZebuAF) and Asian indicine (ZebuAS) groups, avoiding outlier populations based on the multidimensional scaling (MDS) and neighbour-net analyses.



**FIGURE 2** Example of two initial modelled scenarios for determining the domestication history of cattle using approximate Bayesian computation (ABC). Scenario 1 models only two domestication events (black circles) that coincide with the divergence of taurine and indicine cattle. Scenario 4 models three domestication events: one in indicine cattle prior to divergence within the indicine group and two within taurine cattle, after divergence within the taurine group

ABCtoolbox using the 1,000 simulations that produced summary statistics closest to the observed data. The retained simulations were compared to the observed data under a generalized linear model (GLM) to produce a posterior probability  $p$ -value. The  $p$ -value is the proportion of simulations that have a smaller or equal likelihood to the observed data; larger  $p$ -values represent a better fitting GLM and therefore a good match between the simulated and observed scenarios. Different scenarios were compared with Bayes factors (BF), measured as the quotient of the marginal density (MD) of one scenario divided by the MD of an alternative one; a BF greater than 3 is sufficient to reject the alternative scenario (Wegmann et al., 2010). Upon selecting the best-fitting model of the initial four scenarios proposed, alternative scenarios were hypothesized using this as a base model, and the process was repeated multiple times until all 15 scenarios were tested (Supporting Information Figure S1). More complex scenarios were not tested to avoid uninformative overparameterization. Finally, the eight best-fitting scenarios were repeated with the additional two breed sets as replicates.

TreeMix (Pickrell & Pritchard, 2012) was used to address any ambiguity between ABC models differing by only migration. The three replicate breed sets used in the ABC analyses were merged together, and a maximum-likelihood tree was constructed using TreeMix under default parameters. Iteratively, weighted migration edges were added to the network, until 99.8% of the variance of the ancestry between the populations was explained by the model.

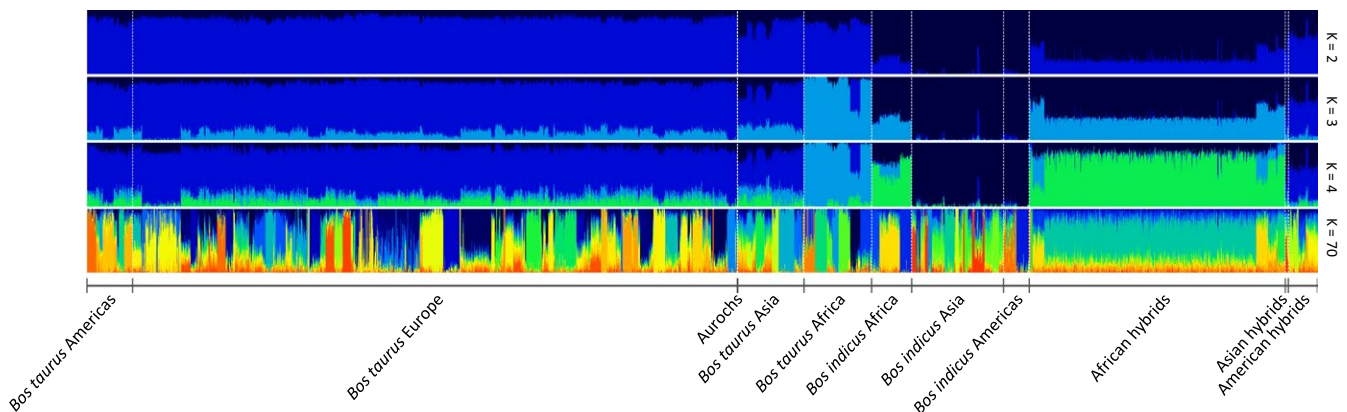
### 3 | RESULTS

#### 3.1 | Admixture and population structure

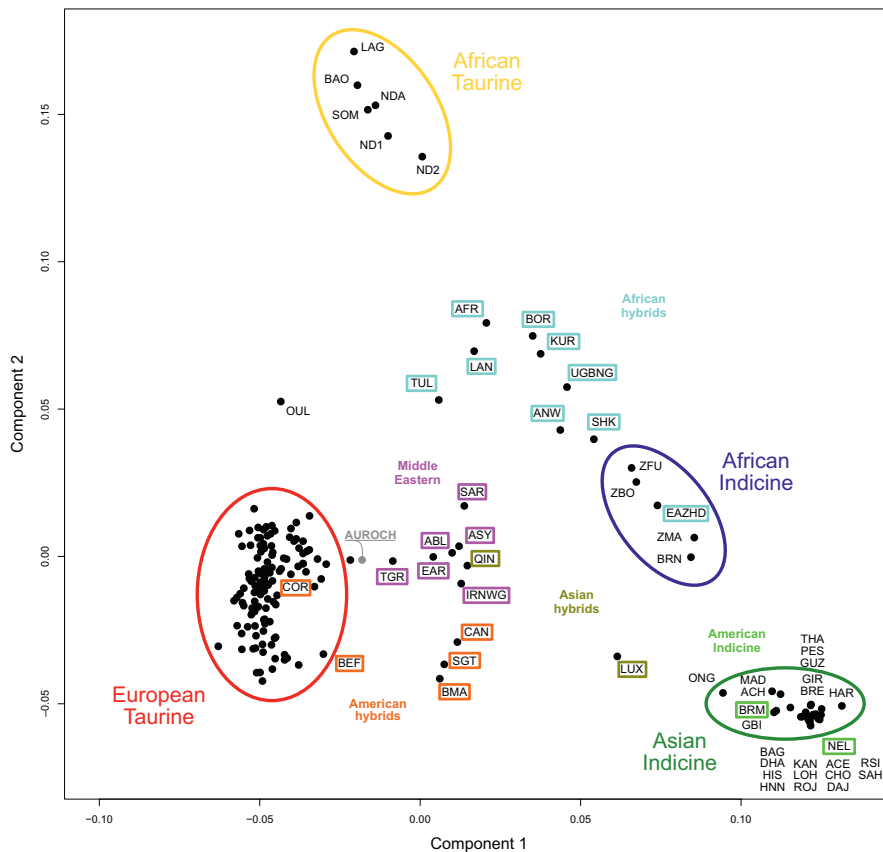
Admixture analysis of 180 populations was performed for 61 different values of  $K$ , ranging between 1 and 150. The CV error across all individuals was lowest at  $K = 70$  (0.545) (Supporting Information Figure S2); however, 56% of the reduction in CV error (CV error difference 0.06) was observed from  $K = 1$  to 5. At  $K = 2$ , the main separation was between *B. taurus* and *B. indicus* (Figure 3, for more

information on the breeds see Supporting Information Figure S3), while for  $K = 3$  the African taurine breeds formed a separate cluster from the European taurine populations (Figure 3).  $K = 4$  showed the split between Asian and African indicine breeds (Figure 3). At higher levels of  $K$ , many taurine breeds could be differentiated from each other, whereas for indicine breeds, the main differentiation was observed for Asiatic breeds, while the African indicine breeds were more similar to each other, presenting a similar level of admixture (data not shown).

Multidimensional scaling analyses resulted in the two major axes explaining ~31% and ~15% of the variance, respectively. The first component separated taurine from indicine cattle, with hybrid breeds occurring in the middle between the two main cattle groups, while the second component separated African taurine from the rest of the cattle breeds (Figure 4). The African hybrids between taurine and indicine breeds were placed between the African taurine and the African indicine populations, while the Asian hybrids were placed between the European taurine and Asiatic indicine cattle. These results were supported by the breed clustering observed in the neighbour-net analysis, which clearly separated Eurasian taurine, African taurine and the indicine breeds. African indicine cattle occurred between the Asiatic indicine populations and the area of the network where the taurine-indicine hybrids were clustered (Figure 5). Overall, both analyses depicted two main clusters of hybrids between *B. taurus* and *B. indicus* (Figures 4 and 5): (a) African hybrids (AFR, ANW, BOR, BRN, EAZHD, KUR, LAN, SHK, TUL, UGBNG), closer to African indicine breeds; and (b) American hybrids (BEF, BMA, CAN, COR, SGT), closer to European taurine cattle. More intermediate positions were assigned to Asian hybrids (LUX, QIN). Apart from hybrids, the Middle Eastern breeds (IRNWG, TGR, EAR, ABL, ASY, SAR) showed a slightly higher proximity to the Eurasian taurine component (Figures 4 and 5). Creole cattle (CCC, CRK, RMS, SNM, SNP, TXL) were placed in the European taurine group, while the American indicine breeds (BRM, NEL) clustered along Asian indicine populations (Figure 5). Finally, the aurochs was placed between the European and Middle Eastern taurine groups.



**FIGURE 3** Admixture individual assignment plots for 179 cattle populations and one auroch sample for  $K = 2, 3, 4$  and 70. Each vertical bar represents an individual, and the proportion of each colour in that bar corresponds to the ancestry (genetic variation) of an individual deriving from a given cluster



**FIGURE 4** Multidimensional scaling (MDS) plot of 3,197 individuals belonging to 180 populations of *Bos primigenius primigenius*, *Bos indicus*, *Bos taurus* and hybrids

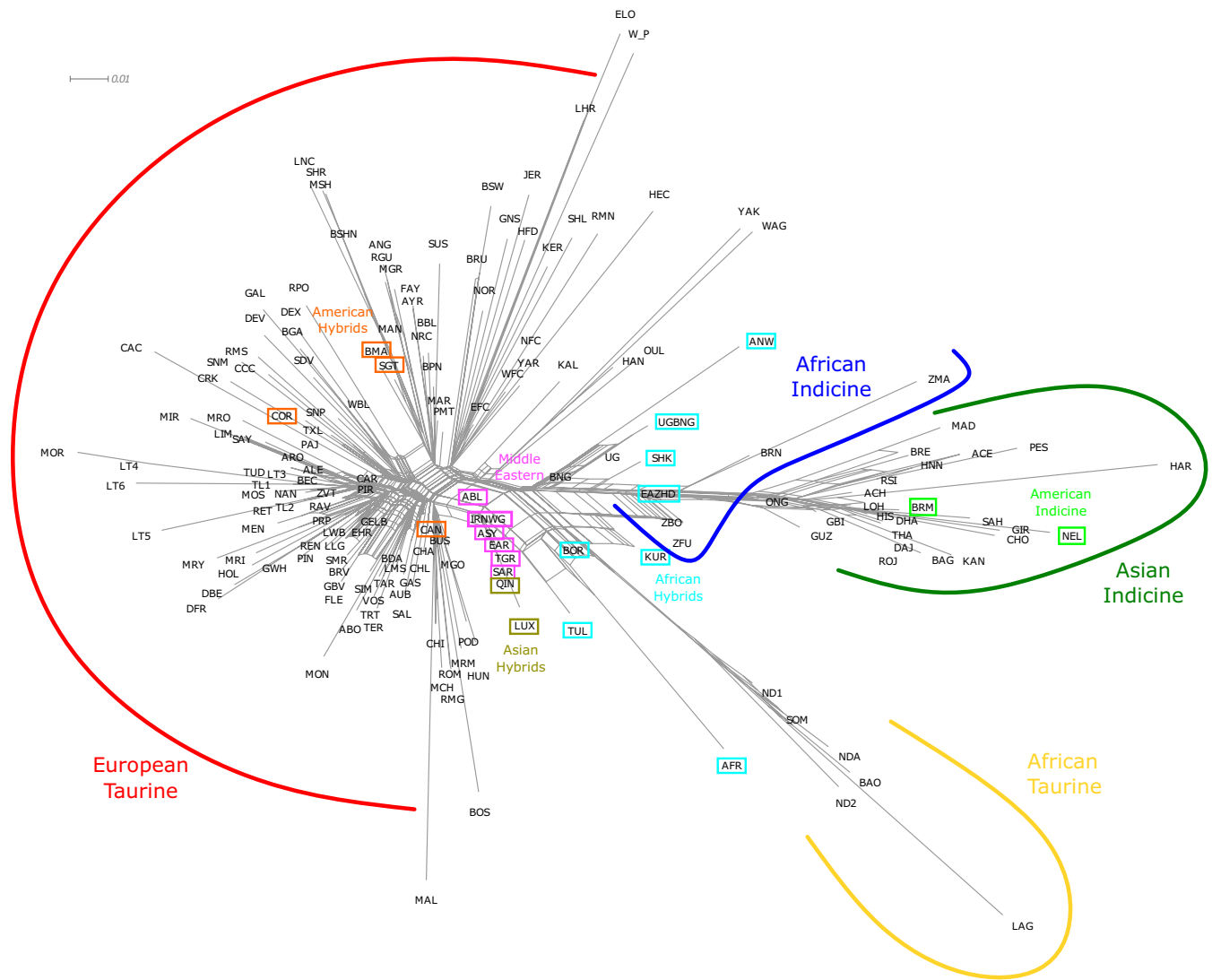
### 3.2 | Approximate Bayesian computation (ABC)

Approximate Bayesian computation was used to reconstruct the demographic history of taurine and indicine breeds. We first assessed whether the reduced data set (2K) resulted in qualitatively the same admixture pattern as the larger data set (8K). The CV error values for K from 1 to 5 were similar for both data sets (less than 1% difference for K = 3 and K = 4; Supporting Information Figure S4). For K = 4, each breed was assigned to one cluster with the ZebuAF showing admixture partially deriving from TaurAF but mostly from ZebuAS (Supporting Information Figure S4), while for K = 3 the clusters corresponded to TaurAF, TaurEU and ZebuAS. Both results reflected the divergence patterns observed in the population structure analyses.

Of the 34 summary statistics originally chosen for the ABC analyses, 17 were removed because of correlations with other statistics, for example, mean number of alleles per locus and pairwise differences between populations due to the high correlation with heterozygosity (Supporting Information Table S3; Figure S5). The remaining 17 summary statistics included measurements of diversity (e.g., mean and standard deviation of heterozygosity across loci for each population) and pairwise divergence (e.g., mean and standard deviation of  $F_{ST}$ ; Supporting Information Table S3). Each observed summary statistics was within the 95% quantiles of the simulated data from that summary statistics, consistently across all scenarios and replicates.

The test of the first four scenarios showed a  $BF > 3$  when comparing scenarios 1 and 2 with two domestication events, over

scenario 4, which included three episodes of domestication (scenario 3 with two domestication events showed the lowest MD and  $p$ -value, reflecting its unsuitability to represent the observed genetic data). Although additional models with three domestication events were tested, scenarios with only two domestications were better supported than those with three domestication events in general (Table 2). Of the 15 modelled scenarios, initial testing showed low support for scenarios 1 to 7 and therefore was not replicated. Among the replicated scenarios, three were identified consistently as best fitting across the three breed sets tested (in bold in Table 2). Scenario 8 displayed high MD and  $p$ -values, with the following features: (a) domestication occurred on two independent occasions and (b) a bidirectional migration event between TaurAF and ZebuAF. The other two scenarios with highest MD and  $p$ -values built on scenario 8 with the addition of a unidirectional (scenario 11) or bidirectional (scenario 14) migration between ZebuAS and TaurEU (Figure 6). The MD difference between these three scenarios was too small to discriminate between them (i.e.,  $BF < 3$ ; Table 2), with estimates of the posterior distributions of parameters such as the effective population size remaining relatively consistent across all three scenarios. These effective population size estimates suggest that following the divergence between the lineages that led the *B. taurus* and *B. indicus* (~250,000 YA) the ancestral taurine and indicine lineages grew demographically from  $\sim 10^3$  until reaching effective population sizes of  $\sim 10^5$  to  $10^6$ ; however, between ~3,600 and 7,900 generations ago these reduced until reaching  $\sim 10^4$  (Supporting Information Table S4). The TreeMix analysis recovered the expected



**FIGURE 5** Neighbour-net using  $F_{ST}$  distances for 174 populations of *Bos indicus*, *Bos taurus* and hybrids featuring more than one sample (see Supporting Information Table S1 for label information). Scale for  $F_{ST}$  distance is displayed in the top left

phylogenetic relationships between the combined breed sets before migration edges were added. Four migration edges were added before the model explained over 99.8% of the variance of the ancestry between the twelve populations (Figure 7). The first two edges were directed from TaurAF to ZebuAF, each weighted between 0.29 and 0.32 of the proportional ancestry received from the TaurAF into ZBO and ZFU. Additionally, a migration edge occurred in the reverse direction between ZebuAF (ZMA) and TaurAF at a similar strength (0.24). Finally, a weak (0.05) migration edge from TaurEU (MON) into the ancestor of ZebuAS was also observed.

## 4 | DISCUSSION

Success in domesticating many species of animals and plants facilitated the gradual change in human societal behaviour, shifting from nomadic foragers and hunters to a sedentary farming society (Diamond, 2002). Cattle provide humans with food, hides and

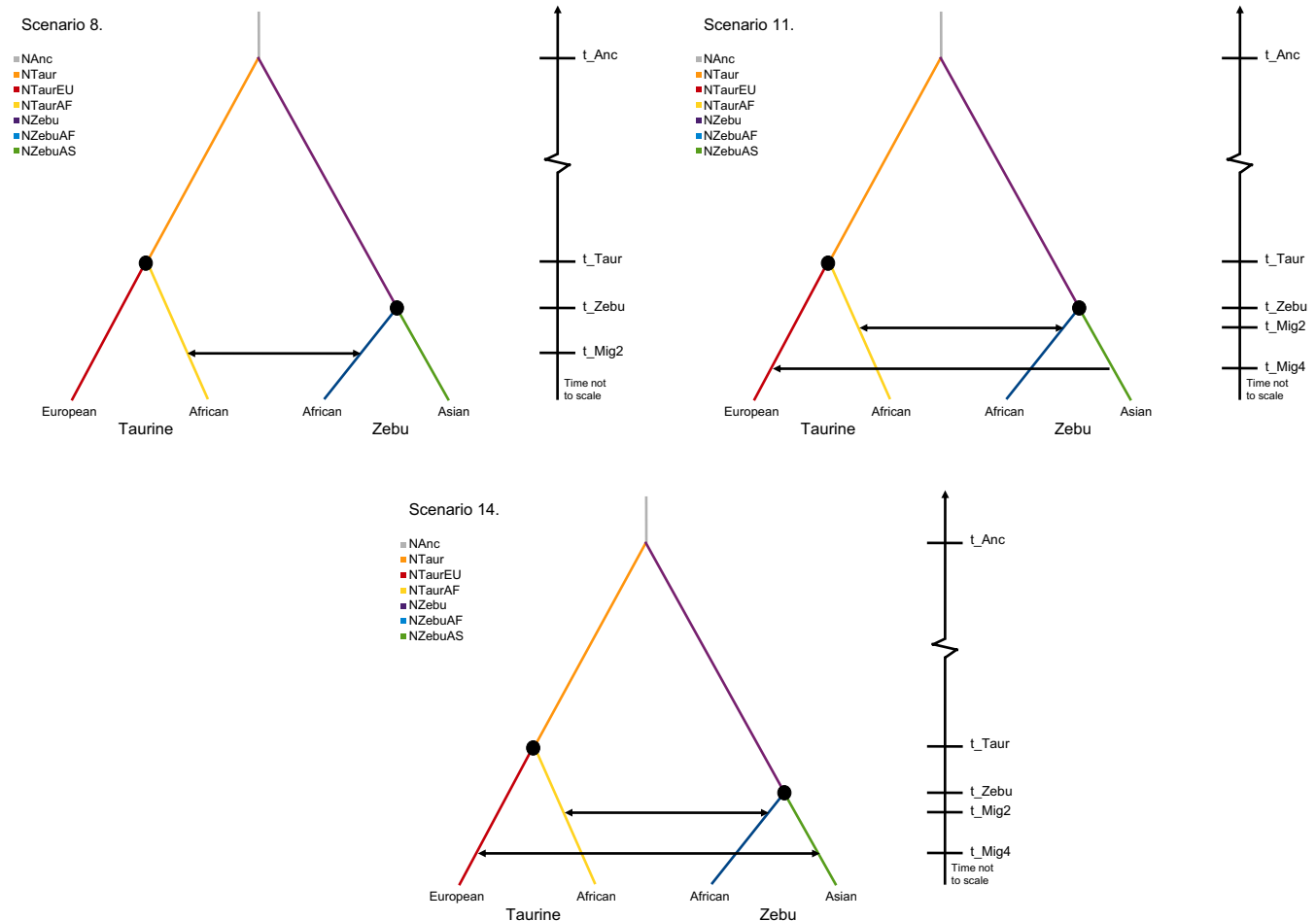
draught power for farming, transport and construction. Animal husbandry and food production may have contributed to the first uneven divisions of labour within early human societies (Ajmone-Marsan et al., 2010). The associated population growth would have led to increased demographic pressure and encouraged dispersal of human populations out of the domestication centres, often accompanied by cattle (Taberlet, Coissac, Pansu, & Pompanon, 2011), helping farmers to colonize habitats that were unavailable to indigenous hunter-gatherers. Our worldwide survey on 180 populations using SNP array data depicted genetic ancestries, admixture, introgression and migration patterns of cattle at a global scale, and gained insights into the long-standing debate on the existence of a third domestication event from aurochs in north-east Africa. However, there are potential issues regarding ascertainment bias associated with the SNP arrays used in this study (Gautier et al., 2010; Matukumalli et al., 2009; Orozco-terWengel et al., 2015), as reflected by the significantly higher genetic variation within taurine cattle when compared to indicine breeds. Higher  $H_o$  in European taurine populations and a



**TABLE 2** Approximate Bayesian computation (ABC) results for different hypothesized domestication histories within taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle

Scenario	Breed set 1 <sup>a</sup>		Breed set 2 <sup>a</sup>		Breed set 3 <sup>a</sup>		Domestication Events	Scenario Description <sup>c</sup>
	MD <sup>b</sup>	<i>p</i> -value <sup>b</sup>	MD <sup>b</sup>	<i>p</i> -value <sup>b</sup>	MD <sup>b</sup>	<i>p</i> -value <sup>b</sup>		
1	13,685.50	0.29					2	<i>B. t.</i> domestication at the time of <i>B. t.</i> divergence, <i>B. i.</i> domestication at the time of <i>B. i.</i> divergence.
2	1,688.92	0.77					2	<i>B. t.</i> domestication at time of <i>B. t.</i> divergence, <i>B. i.</i> domestication before <i>B. i.</i> divergence.
3	22.14	0.15					2	Domestications each before divergence of <i>B. t.</i> and <i>B. i.</i>
4	325.70	0.73					3	<i>B. t.</i> domestications after <i>B. t.</i> divergence, <i>B. i.</i> domestication before <i>B. i.</i> divergence.
5	0.51	0.08					3	Scenario 4 with ancient <i>B. t.</i> divergence (lower bound at 12,500 generations)
6	1,059.53	0.97					3	<i>B. t.</i> domestications after <i>B. t.</i> divergence, <i>B. i.</i> domestication at the time of <i>B. i.</i> divergence.
7	5,747.75	0.26					2	Scenario 1 with bidirectional migration before either divergence.
8	<b>578,215.00</b>	<b>0.82</b>	<b>531,478.00</b>	<b>0.66</b>	<b>237,680.00</b>	<b>0.46</b>	2	<b>Scenario 1 with bidirectional migration between African <i>B. t.</i> and African <i>B. i.</i> after both divergences.</b>
9	329,323.00	0.77	991,43.70	0.58	100,88.40	0.61	2	Scenario 8 with bidirectional migrations before either divergence.
10	244,663.00	0.77	273,67.20	0.55	321,27.50	0.49	2	Scenario 8 with unidirectional migration from African <i>B. i.</i> to European <i>B. t.</i> after both divergences.
11	<b>360,715.00</b>	<b>0.86</b>	<b>592,151.00</b>	<b>0.75</b>	<b>491,523.00</b>	<b>0.70</b>	2	<b>Scenario 8 with unidirectional migration from Asian <i>B. i.</i> to European <i>B. t.</i> after both divergences.</b>
12	210,936.00	0.78	327,979.00	0.62	231,650.00	0.50	2	Scenario 8 with unidirectional migration from African <i>B. i.</i> and Asian <i>B. i.</i> to European <i>B. t.</i> after both divergences.
13	2.37	0.00	598.91	0.14	22.44	0.06	2	Scenario 8 with constant ongoing bidirectional migration between <i>B. t.</i> and <i>B. i.</i> before either divergence
14	<b>194,460.00</b>	<b>0.83</b>	<b>673,869.00</b>	<b>0.84</b>	<b>620,562.00</b>	<b>0.72</b>	2	<b>Scenario 8 with bidirectional migration between Asian <i>B. i.</i> and European <i>B. t.</i> after both divergences.</b>
15	124,482.00	0.66	235,974.00	0.68	364,806.00	0.58	2	Scenario 14 with ancient bottlenecks within <i>B. t.</i> and <i>B. i.</i>

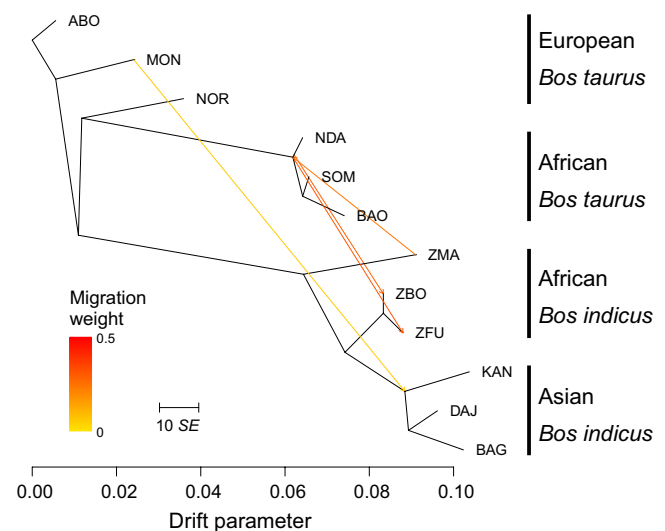
*Note.* The three best scenarios with consistently high MD values across breed sets are highlighted in bold. Models are shown in Figures 2,6 and Supporting Information Figure S1.  
<sup>a</sup>Breed set 1: Normande (NOR), Somba (SOM), Zebu from Madagascar (ZMA), Kankraj (KAN); Breed set 2: Abondance (ABO), N'dama (NDA), Zebu Fulani (ZFU), Bhagnari (BAG); Breed set 3: Montbeliard (MON), Baoule (BAO), Zebu Bororo (ZBO), Dajal (DAJ). <sup>b</sup>Marginal density (MD) and *p*-values are taken from the 1,000 simulations most similar to the observed data from 1 million simulations. <sup>c</sup>Divergences here refer to events within *Bos taurus* (*B. t.*) and *Bos indicus* (*B. i.*) rather than the ancient divergence between them from *Bos primigenius primigenius*.



**FIGURE 6** The three best-modelled scenarios for the domestication history of cattle using approximate Bayesian computation (ABC). Domestication episodes and migratory events between populations are shown by black circles and arrows, respectively

potential overestimation of the inbreeding coefficient ( $F_{IS}$ ) in indicine cattle would be an unexpected result had the data set used an array not biased towards *B. taurus* variation (Supporting Information Table S1). Ascertainment bias was reduced through removing a proportion of the markers in LD, and highly correlated markers with similar genealogical history were removed from the data set, reducing multicollinearity effects that would lead to overestimation of differentiation measurements. LD pruning of SNPs has been shown to be effective at reducing heterozygosity overestimations when compared to whole-genome sequencing and, although this methodology underestimates  $F_{ST}$ , it is consistent regardless of relatedness to the ascertainment populations (Malomane et al., 2018). Furthermore, by replicating the ABC scenarios to increase the sampling scheme, we attempted to reduce any breed specific biases and more effective at capturing the demographic history representative of whole subgroups (Rougemont et al., 2016).

The Admixture analysis best supported the partition for  $K = 70$ ; however, the low variation in CV error for  $K$  values between 15 and 100 suggests that sample groupings above  $K = 15$  only reflect minor improvements in the resolution of the clusters detected. This observation is likely to be an outcome of several factors influencing the statistical power to discriminate between the populations analysed,



**FIGURE 7** Phylogenetic network of the inferred relationships between 12 cattle breeds estimated using TreeMix. Migration edges between breeds are shown as arrows pointing towards the recipient population and coloured according to the proportional ancestry received from the donor population. Scale bar is 10 times the mean standard error of the estimated entries in the covariance matrix

for example, the presence of known hybrids in the data set, low sample sizes for some breeds ( $n < 8$ ), or the presence of related individuals (e.g., family structure) (Barbato et al., 2017; Kijas et al., 2012; McTavish et al., 2013; Orozco-terWengel et al., 2015). This effect may be further exacerbated due to shared genetic variation and incomplete lineage sorting among populations, reflecting their low divergence since they were founded ~200 YA (Bradley et al., 1996). An additional factor influencing the power to separate populations is the presence of gene flow between them, as reflected by the admixture signals in African indicine cattle. Deriving from Asiatic indicine populations introduced to Africa ~2,500–3,500 YA, African indicine cattle represent an important component of African taurine genetic diversity, probably reflecting gene flow between these populations due to their geographic proximity, as well as deliberate hybridization by farmers to incorporate local adaptations (e.g., trypanotolerance) into the indicine animals following their entry into Africa (Troy et al., 2001). Analyses with Admixture for low  $K$  values initially divided taurine and indicine cattle ( $K = 2$ ; Figure 3), as expected due to their relatively ancient divergence from their common ancestor, *B. p. primigenius*. Increasing  $K$  to 3 and 4 caused the within-species division of taurine into African and Eurasian and of indicine into African and Asian genomic components, respectively (Figure 3). Interestingly, for low clusters ( $K = 2$ –4), African hybrids display similar proportional ancestry to African indicine. The introgression of indicine into African taurine cattle is predicted to have initially occurred ~4,000 YA, resulting in a consistently high indicine component within African hybrid cattle (Bahbahani et al., 2017), possibly reinforced by modern admixture with African indicine.

The neighbour-net and MDS analyses identified similar patterns of divergence between indicine and taurine populations (Figures 4 and 5). The first two components of variation in the MDS analysis also identified the taurine vs. indicine (~31%) and African vs. Eurasian taurine (~15%) splits. While the division between the taurine groups explained much less genetic variance than the first component, it still explained a substantial proportion of the total (and more than twofold the third component), suggesting substantial differentiation between the two taurine groups, consistent with the hypothesis that the two groups harbour rather distinct genetic pools. In contrast to this, African indicine populations strongly resembled Asiatic indicine breeds. The relatively short distances between African indicine to both Asian indicine and African taurine populations in the neighbour-net suggest admixture between these groups. Furthermore, the central block of the neighbour-net depicts multiple complex connections among breeds, suggesting relatively recent divergence across many populations (Feliuss, 1995). Despite being classified as *B. taurus*, Middle Eastern breeds occupy central positions both in the neighbour-net and MDS analyses, suggesting an admixed genetic background with an indicine component, as has been observed for several breeds from the Middle East and Fertile Crescent (Decker et al., 2014; Karimi et al., 2016; Loftus et al., 1999). Consistent with the hypothesis of admixture of European taurine populations with local aurochs prior to their extinction approximately 400 YA (Achilli et al., 2008; Park et al., 2015; Upadhyay

et al., 2017), the *B. p. primigenius* sample in this study is placed close to Eurasian taurine groups in the MDS analysis. Although Decker et al. (2014) suggested a significant admixed aurochs ancestry for African taurine populations, with aurochs contributing up to 26% of their genomes, our results failed to detect this influence, almost certainly because the British aurochs used is not an appropriate proxy for African aurochs—often classified as *B. p. africanus* or *B. p. opisthonotus* (Clutton-Brock, 1999). Similar to the Admixture results (Figure 3), African hybrids from *B. taurus* and *B. indicus*, that were expected to be found in the central areas of the neighbour-net and MDS graphics, were placed near to the African indicine group, pointing towards a higher contribution of the African indicine genepool in the formation of those hybrid breeds, as opposed to American hybrids that displayed higher influences from European taurine breeds. The genetic composition of Asian hybrids (QIN, LUX) was more balanced between Asian indicine and taurine origins. As expected, Creole cattle were placed closer to the European taurine cluster (Decker et al., 2014; Martínez et al., 2012; Pitt et al., 2018), and American indicine cattle clustered with Asian indicine breeds (Orozco-terWengel et al., 2015).

The substantial differentiation between the European and African taurine populations when compared to African and Asian indicine groups may be explained by several factors, such as lower gene flow between Africa and Europe than between Africa and Asia or a more recent divergence of indicine cattle. Another explanation may be a stronger effect of genetic drift in African taurine populations with respect to Eurasian taurine or indicine populations. Genetic drift would increase the rate of the change in allele frequencies in populations with a smaller effective population size ( $N_e$ ), increasing the probability of accumulating differences between populations. An analysis of the trajectories of the effective population size over the last ~8,000 years shows that most cattle populations around the world exhibit relatively large effective population sizes at the beginning of the Holocene (~3,000  $N_e$ ); however, more recently, all cattle populations have experienced a drastic decrease in  $N_e$  reaching modern  $N_e$  values of ~500 or less for many populations (Barbato, Orozco-Terwengel, Tapio, & Bruford, 2015; Orozco-terWengel et al., 2015), with the African taurine population showing, on average, larger  $N_e$  values than Eurasian taurine populations, albeit only marginally.

The hypothesis of an additional domestication event in Africa from local aurochs has been also proposed to explain the large difference between taurine branches. We explicitly tested the hypothesis of whether three domestication events better explain the genetic variation observed in modern domestic cattle when compared to the more widely accepted two domestication scenario. ABC tested on a number of demographic scenarios showed reasonable support for three domestication events (scenarios 4 and 6) in terms of how well the estimated GLM fitted the observed data (i.e., the  $p$ -value). However, when scenarios including three domestications were compared to similar models in which only two domestications occurred, the latter received stronger support ( $BF > 3$ ) (Table 2). The rejection of a hypothesized third domestication suggests an alternative

explanation for the differences observed between African and Eurasian taurine populations: admixture between migrating domestic cattle from the Middle East and indigenous African. Domestic cattle and their wild relatives occupied the same geographic regions for a long period of time, which raises the possibility that both taurine and indicine cattle naturally hybridized with aurochs or that local farmers mixed them with local aurochs to restock their herds (McTavish et al., 2013; Troy et al., 2001). Uniparental loci such as mtDNA and Y-chromosome studies have generally underplayed the significance of admixture with wild aurochs, as haplotypes present in ancient DNA samples are often closely related but phylogenetically distinct from those in extant cattle samples (Edwards et al., 2010; Götherström et al., 2005; Schibler, Elsner, & Schlumbaum, 2014; Troy et al., 2001). However, there is evidence of gene flow from wild aurochs prior to the extinction (~400 YA) into extant cattle in areas such as Italy, Iberia, southern Europe and the British Isles (Achilli et al., 2008; Park et al., 2015; Upadhyay et al., 2017). These results are consistent with those in other domesticated animals such as horses (Warmuth et al., 2012), pigs (Frantz et al., 2015) and camels (Almathen et al., 2016), where gene flow with wild populations has been observed in modern breeds.

Previous studies using SNP data genotyped in taurine and indicine cattle have found an African taurine component in European taurine cattle reaching an average of 10% of their genotype, while between 5% and 10% of the genetic background in some European breeds is of indicine origin (e.g., Decker et al., 2014; McTavish et al., 2013). We used ABC to test alternative migration hypotheses between the sets of African and Eurasian taurine and indicine breeds used to model the domestication events. However, it was not possible to differentiate between a scenario with bidirectional gene flow between African breeds, and two alternative ones that further included unidirectional or bidirectional gene flow between the Eurasian breeds. Further specification of the model was unsuccessful and trying to define the presence or manner of migration between other groups only yielded insignificant changes between models. This could be due to the difficulty in accurately reconstructing demographic history and model differentiation, for example, divergence and gene flow can be indistinguishable to isolation followed by secondary contact, or brief bottlenecks remaining unidentified if preceded by large expansions due to the increased genetic diversity (Gray et al., 2014; Rougemont et al., 2016). These limitations restrict the intrinsic complexity in the models, ideally, events such as divergences and domestication would be reconstructed as gradual processes occurring over many generations as cattle were shifted from prey animals, to managed herds, to captive bred stock (Larson & Burger, 2013). Despite efforts to limit overparameterization by reducing the ranges of many parameters, both the number of prior parameters and their range may have reduced statistical power of posterior estimates (Wegmann et al., 2010). The progression of refining the ABC model through serial selection and diversification of preferentially chosen scenarios was the most computationally efficient methodology to define increasingly more complex models. “Less

favourable” models were rejected at each step to constrain the otherwise very large number of models simulated. Although this saves significant time and resources, it is built on the assumption that discarded models will never be viable regardless of alterations; unfortunately, this compromise, along with the reduction in the data set from 8K to 2K, was necessary logistical restrictions. However, the first three migration edges incorporated into the maximum-likelihood tree generated in TreeMix recaptured the same migratory patterns between TaurAF and ZebuAF as tested in ABC scenario 8 (Figures 6 and 7) and explained up to 99.78% of the variance in the tree. Adding a fourth migration edge marginally increased the explanation of the variance in the tree and corresponded to a weak migration edge from TaurEU into ZebuAS. Overall, this supports the ABC results and indicates that the more important gene flow between populations occurred between *B. taurus* and *B. indicus* within Africa, although minor movements of cattle between Europe and Asia may also contribute to shape the population structure we see today.

In conclusion, SNP array data collected in more than three thousand cattle samples belonging to 180 populations with a worldwide geographic distribution encompassing Africa, Europe, Asia and the Americas continents provided a comprehensive picture of genetic diversity, population structure and demographic dynamics of cattle populations at a global level. The analyses confirm the large differentiation between African and Eurasian taurine and the high levels of admixture in African indicine cattle from both Asian indicine and African taurine cattle, and reveal a higher contribution from African indicine genetic origin in the formation of African hybrids, as opposed to American hybrids which exhibited a higher influence from taurine breeds. Modelling the domestication history of cattle using approximate Bayesian computation consistently favoured scenarios involving only two domestication events, discarding a third *B. p. primigenius* domestication in Egypt and suggesting the subsequent hybridization from local aurochs to explain the additional genetic variation detected. Further analysis exploring more domestication situations, particularly focussed on migration between groups (e.g., African indicine and European taurine cattle), will help to disentangle the complex human-mediated microevolution of domestic cattle, in particular if using data deriving from whole-genome sequence data that suffer less of ascertainment bias as SNP arrays do. Paleogenomics analyses of Middle Eastern and African wild aurochs predating domestication and early Middle Eastern and African domestic cattle will provide the data required to fully address these questions.

## ACKNOWLEDGEMENTS

We thank Prof. J. Cañón and Prof. S. Dunner from Universidad Complutense de Madrid (UCM) for their contribution of SNP array data. NS is a recipient of a Marie Skłodowska-Curie Fellowship and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No DLV-655100. DP is funded by BBSRC SWBio DTP PhD studentship (BB/



M009122/1). MWB and POTW are funded by BBSRC through FACCE-JPI ERA-NET Smart Agriculture Program (CLIMGEN CONSORTIUM; BB/M019276/1).

## CONFLICT OF INTEREST

None declared.

## ORCID

Daniel Pitt  <http://orcid.org/0000-0002-8429-2957>

Natalia Sevan  <http://orcid.org/0000-0003-4766-6291>

Michael W. Bruford  <http://orcid.org/0000-0001-6357-6080>

Pablo Orozco-terWengel  <http://orcid.org/0000-0002-7951-4148>

## REFERENCES

- Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N., ... Battaglia, V. (2008). Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Current Biology*, 18(4), R157–R158. <https://doi.org/10.1016/j.cub.2008.01.019>
- Ajmone-Marsan, P., Garcia, J. F., & Lenstra, J. A. (2010). On the origin of cattle: How aurochs became cattle and colonized the world. *Evolutionary Anthropology: Issues, News, and Reviews*, 19(4), 148–157. <https://doi.org/10.1002/evan.20267>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Almathen, F., Charreau, P., Mohandesan, E., Mwacharo, J. M., Orozco-terWengel, P., Pitt, D., ... Burger, P. (2016). Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. *Proceedings of the National Academy of Sciences*, 113(24), 6707–6712. <https://doi.org/10.1073/pnas.1519508113>
- Applegate, A., Gautier, A., & Duncan, S. (2001). The north tumuli of the Nabta Late Neolithic ceremonial complex. In F. Wendorf & R. Schild, & Associates (Eds.), *Settlement of the Egyptian Sahara. Volume 1: The Archaeology of Nabta Playa* (pp. 468–488). New York, NY: Kluwer. <https://doi.org/10.1007/978-1-4615-0653-9>
- Bahbahani, H., Tijjani, A., Mukasa, C., Wragg, D., Almathen, F., Nash, O., ... Hanotte, O. (2017). Signatures of selection for environmental adaptation and zebu × taurine hybrid fitness in East African Shorthorn Zebu. *Frontiers in Genetics*, 8, 1–20. <https://doi.org/10.3389/fgene.2017.00068>
- Barbato, M., Hailer, F., Orozco-terWengel, P., Kijas, J., Mereu, P., Cabras, P., ... Bruford, M. W. (2017). Genomic signatures of adaptive introgression from European mouflon into domestic sheep. *Scientific Reports*, 7(1), 7623. <https://doi.org/10.1038/s41598-017-07382-7>
- Barbato, M., Orozco-Terwengel, P., Tapio, M., & Bruford, M. (2015). SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, 6, 109. <https://doi.org/10.3389/fgene.2015.00109>
- Beja-Pereira, A., Caramelli, D., Lalueza-Fox, C., Vernesi, C., Ferrand, N., Casoli, A., ... Martini, A. (2006). The origin of European cattle: Evidence from modern and ancient DNA. *Proceedings of the National Academy of Sciences*, 103(21), 8113–8118. <https://doi.org/10.1073/pnas.0509210103>
- Bonfiglio, S., Jinja, C., De Gaetano, A., Achilli, A., Olivieri, A., Colli, L., ... Ferretti, L. (2012). Origin and spread of *Bos taurus*: New clues from mitochondrial genomes belonging to haplogroup T1. *PLoS ONE*, 7(6), e38601. <https://doi.org/10.1371/journal.pone.0038601>
- Bovine Hapmap Consortium, Gibbs, R. A., Taylor, J. F., Van Tassell, C. P., Barendse, W., Eversole, K. A., ... Dodds, K. G. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, 324, 528–532. <https://doi.org/10.1126/science.1167936>
- Bradley, D. G., MacHugh, D. E., Cunningham, P., & Loftus, R. T. (1996). Mitochondrial diversity and the origins of African and European cattle. *Proceedings of the National Academy of Sciences*, 93(10), 5131–5135. <https://doi.org/10.1073/pnas.93.10.5131>
- Bruford, M. W., Bradley, D. G., & Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics*, 4(11), 900–910. <https://doi.org/10.1038/nrg1203>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Clutton-Brock, J. (1999). *A Natural History of Domesticated Mammals*. Cambridge, UK: Cambridge University Press.
- Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J., Alcalá, A. M., Sonstegard, T. S., ... Taylor, J. F. (2014). Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics*, 10(3), e1004254. <https://doi.org/10.1371/journal.pgen.1004254>
- Decker, J. E., Pires, J. C., Conant, G. C., McKay, S. D., Heaton, M. P., Chen, K., ... Taylor, J. F. (2009). Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proceedings of the National Academy of Sciences*, 106(44), 18644–18649. <https://doi.org/10.1073/pnas.0904691106>
- Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature*, 418(6898), 700–707. <https://doi.org/10.1038/nature01019>
- Edwards, C. J., Magee, D. A., Park, S. D., McGettigan, P. A., Lohan, A. J., Murphy, A., ... Bradley, D. G. (2010). A complete mitochondrial genome sequence from a mesolithic wild aurochs (*Bos primigenius*). *PLoS ONE*, 5(2), e9255. <https://doi.org/10.1371/journal.pone.0009255>
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, 9(10), e1003905. <https://doi.org/10.1371/journal.pgen.1003905>
- Excoffier, L., & Foll, M. (2011). Fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27(9), 1332–1334. <https://doi.org/10.1093/bioinformatics/btr124>
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Felius, M. (1995). *Cattle Breeds: An Encyclopedia* (ed. 1). Doetinchem, The Netherlands: C Misset bv.
- Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27–32. <https://doi.org/10.1111/1755-0998.12509>
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H. J., Cagan, A., Bosse, M., ... Groenen, M. A. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nature Genetics*, 47(10), 1141–1148. <https://doi.org/10.1038/ng.3394>
- Gautier, M., Laloë, D., & Moazami-Goudarzi, K. (2010). Insights into the genetic history of French cattle from dense SNP data on 47 worldwide breeds. *PLoS ONE*, 5(9), e13038. <https://doi.org/10.1371/journal.pone.0013038>
- Götherström, A., Anderung, C., Hellborg, L., Elburg, R., Smith, C., Bradley, D. G., & Ellegren, H. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1579), 2345–2351. <https://doi.org/10.1098/rspb.2005.3243>

- Gray, M. M., Wegmann, D., Haas, R. J., White, M. A., Gabriel, S. I., Searle, J. B., ... Payseur, B. A. (2014). Demographic history of a recent invasion of house mice on the isolated Island of Gough. *Molecular Ecology*, 23(8), 1923–1939. <https://doi.org/10.1111/mec.12715>
- Grigson, C. (1991). An African origin for African cattle? Some archaeological evidence *African Archaeological Review*, 9, 119–144. <https://doi.org/10.1007/BF01117218>
- Grigson, C. (2000). *Bos africanus* (Brehm)? Notes on the archaeozoology of the native cattle of Africa. In R. M. Blench & K. C. MacDonald (Eds.), *The Origins and Development of African Livestock. Archaeology, Genetics, Linguistics, and Ethnography* (pp. 38–60). London, UK: UCL Press.
- Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., & Rege, J. E. O. (2002). African pastoralism: Genetic imprints of origins and migrations. *Science*, 296, 336–339. <https://doi.org/10.1126/science.1069878>
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267. <https://doi.org/10.1093/molbev/msj030>
- Iso-Touru, T., Tapio, M., Vilkkii, J., Kiseleva, T., Ammosov, I., Ivanova, Z., ... Kantanen, J. (2016). Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. *Animal Genetics*, 47(6), 647–657. <https://doi.org/10.1111/age.12473>
- Karimi, K., Strucken, E. M., Moghaddar, N., Ferdosi, M. H., Esmailzadeh, A., & Gondro, C. (2016). Local and global patterns of admixture and population structure in Iranian native cattle. *BMC Genetics*, 17(1), 108. <https://doi.org/10.1186/s12863-016-0416-z>
- Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., ... Dalrymple, B. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, 10(2), e1001258. <https://doi.org/10.1371/journal.pbio.1001258>
- Larson, G., & Burger, J. (2013). A population genetics view of animal domestication. *Trends in Genetics*, 29(4), 197–205. <https://doi.org/10.1016/j.tig.2013.01.003>
- Loftus, R. T., Ertugrul, O., Harba, A. H., El-Barody, M. A., MacHugh, D. E., Park, S. D., & Bradley, D. G. (1999). A microsatellite survey of cattle from a centre of origin: The Near East. *Molecular Ecology*, 8, 2015–2022. <https://doi.org/10.1046/j.1365-294x.1999.00805.x>
- Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M., & Cunningham, P. (1994). Evidence for two independent domestications of cattle. *Proceedings of the National Academy of Sciences U S A*, 91, 2757–2761. <https://doi.org/10.1073/pnas.91.7.2757>
- MacHugh, D. E., Larson, G., & Orlando, L. (2017). Taming the Past: Ancient DNA and the Study of Animal Domestication. *Annual Review of Animal Biosciences*, 5, 329–351. <https://doi.org/10.1146/annurev-animal-022516-022747>
- MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P., & Bradley, D. G. (1997). Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics*, 146(3), 1071–1086.
- Malomane, D. K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A. R., & Simianer, H. (2018). Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics*, 19(1), 22. <https://doi.org/10.1186/s12864-017-4416-9>
- Martínez, A. M., Gama, L. T., Cañón, J., Ginja, C., Delgado, J. V., Dunner, S., ... Zaragoza, P. (2012). Genetic footprints of Iberian cattle in America 500 years after the arrival of Columbus. *PLoS ONE*, 7, e49066. <https://doi.org/10.1371/journal.pone.0049066>
- Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., ... Van Tassell, C. P. (2009). Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE*, 4(4), e5350. <https://doi.org/10.1371/journal.pone.0005350>
- Mbole-Kariuki, M. N., Sonstegard, T., Orth, A., Thumbi, S. M., de C Bronsvort, B. M., Kiara, H., & Woolhouse, M. E. J. (2014). Genome-wide analysis reveals the ancient and recent admixture history of East African Shorthorn Zebu from Western Kenya. *Heredity*, 113(4), 297–305. <https://doi.org/10.1038/hdy.2014.31>
- McTavish, E. J., Decker, J. E., Schnabel, R. D., Taylor, J. F., & Hillis, D. M. (2013). New World cattle show ancestry from multiple independent domestication events. *Proceedings of the National Academy of Sciences*, 110(15), E1398–E1406. <https://doi.org/10.1073/pnas.1303367110>
- Orozco-terWengel, P., Barbato, M., Nicolazzi, E., Biscarini, F., Milanese, M., Davies, W., ... Bruford, M. W. (2015). Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Frontiers in Genetics*, 6, 191. <https://doi.org/10.3389/fgene.2015.00191>
- Park, S. D., Magee, D. A., McGettigan, P. A., Teasdale, M. D., Edwards, C. J., Lohan, A. J., ... Chamberlain, A. T. (2015). Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. *Genome Biology*, 16(1), 171. <https://doi.org/10.1186/s13059-015-0790-2>
- Payne, W., & Hodges, J. (1997). *Tropical Cattle, Origins, Breeds and Breeding Policies*. Oxford: Blackwell Science Ltd.
- Pellecchia, M., Negrini, R., Colli, L., Patrini, M., Milanese, E., Achilli, A., ... Ajmone-Marsan, P. (2007). The mystery of Etruscan origins: Novel clues from *Bos taurus* mitochondrial DNA. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1614), 1175–1179. <https://doi.org/10.1098/rspb.2006.0258>
- Pérez-Pardal, L., Royo, L. J., Beja-Pereira, A., Chen, S., Cantet, R. J., Traoré, A., ... Goyache, F. (2010). Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites. *Heredity*, 105(6), 511–519. <https://doi.org/10.1038/hdy.2010.30>
- Pérez-Pardal, L., Royo, L. J., Beja-Pereira, A., Curik, I., Traoré, A., Fernández, I., ... Goyache, F. (2009). Y-specific microsatellites reveal an African subfamily in taurine (*Bos taurus*) cattle. *Animal Genetics*, 41(3), 232–241.
- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8(11), e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
- Pitt, D., Bruford, M. W., Barbato, M., Orozco-terWengel, P., Martinez, R., & Seavane, N. (2018). Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. *Evolutionary Applications*, 58(9), 238. Advance online publication. <https://doi.org/10.1111/eva.12641>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <http://www.R-project.org/>.
- Rougemont, Q., Roux, C., Neuenschwander, S., Goudet, J., Launey, S., & Evanno, G. (2016). Reconstructing the demographic history of divergence between European river and brook lampreys using approximate Bayesian computations. *PeerJ*, 4, e1910. <https://doi.org/10.7717/peerj.1910>
- Schibler, J., Elsner, J., & Schlumbaum, A. (2014). Incorporation of aurochs into a cattle herd in Neolithic Europe: Single event or breeding? *Scientific Reports*, 4(1), 5798.
- Stock, F., & Gifford-Gonzalez, D. (2013). Genetics and African Cattle Domestication. *African Archaeological Review*, 30, 51–72. <https://doi.org/10.1007/s10437-013-9131-6>
- Taberlet, P., Coissac, E., Pansu, J., & Pompanon, F. (2011). Conservation genetics of cattle, sheep, and goats. *Comptes Rendus Biologies*, 334(3), 247–254. <https://doi.org/10.1016/j.crv.2010.12.007>

- Troy, C. S., MacHugh, D. E., Bailey, J. F., Magee, D. A., Loftus, R. T., Cunningham, P., ... Bradley, D. G. (2001). Genetic evidence for Near-Eastern origins of European cattle. *Nature*, 410(6832), 1088–1091. <https://doi.org/10.1038/35074088>
- Upadhyay, M. R., Chen, W., Lenstra, J. A., Goderie, C. R., MacHugh, D. E., Park, S. D., ... Groenen, M. A. (2017). Genetic origin, admixture and population history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity*, 118(2), 169–176. <https://doi.org/10.1038/hdy.2016.79>
- Warmuth, V., Eriksson, A., Bower, M. A., Barker, G., Barrett, E., Hanks, B. K., ... Soyonov, V. (2012). Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proceedings of the National Academy of Sciences*, 109(21), 8202–8206. <https://doi.org/10.1073/pnas.1111122109>
- Wegmann, D., Leuenberger, C., Neuenschwander, S., & Excoffier, L. (2010). ABCtoolbox: A versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, 11(1), 116. <https://doi.org/10.1186/1471-2105-11-116>

Zeuner, F. E. (1963). *A History of Domesticated Animals*. London, UK: Hutchinson.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Pitt D, Sevane N, Nicolazzi EL, et al. Domestication of cattle: Two or three events? *Evol Appl*. 2018;00:1–14. <https://doi.org/10.1111/eva.12674>